Reconfigurable Writing Architecture for Reliable RRAM Operation in Wide Temperature Ranges

Fernando García-Redondo, Hernán Aparicio, Pablo Ituero, Pablo Royer, Marisa López-Vallejo and Carlos López-Barrio

Abstract—**Resistive switching memories (RRAM) are an attractive alternative to non-volatile storage and non-conventional computing systems, but their behavior strongly depends on the cell features, driver circuit and working conditions. In particular, the circuit temperature and the writing voltage scheme become critical issues, determining resistive switching memories performance. These dependencies usually force a design time trade-off among reliability, device endurance and power consumption, and therefore imposing non-flexible functioning schemes and limiting the system performance. In this paper we present a writing architecture that ensures the correct operation no matter the working temperature, and allows the dynamic load of application oriented writing profiles. Thus, taking advantage of more efficient configurations, the system can be dynamically adapted to overcome RRAM intrinsic challenges. Several profiles are analyzed regarding power consumption, temperature-variations protection and operation speed, showing speed-ups near to $700\times$ compared against other published drivers.**

Index Terms—**RRAM, ReRAM, Memristor, Temperature, Reliability, Reconfigurable, Dynamic Writing Driver**

I. Introduction

Memristor based memories, RRAM or ReRAM, are one of the most promising alternatives to future storage and neuromorphic computing systems. Basically, these devices are characterized by their resistive switching capabilities under external stimuli. The resistive switching mechanism is generally based on the formation and disruption of a *conductive filament (CF)* across an insulator/semiconductor material, matching a *Low Resistive State (LRS)* and a *High Resistive State (HRS)*.

However as many emerging technologies, RRAM must solve several integration obstacles before being established as the next standard. Performance and reliability are two key factors that can determine the success of a novel technology over another. Process variations, crossbar design and non-linear selector devices inclusion are the most concerning issues that nowadays limit scope to resistive-switching technologies. To overcome the related design challenges, reliable [1], [2], fast [3], and low-power [1], [3]–[5] writing schemes have been proposed. Although most efforts focus on device to device and cycle to cycle cell variations, temperature is also a critical issue to be addressed in order to ensure the correct behavior of RRAM circuits. The wide temperature ranges under which resistive switching devices operate constrict the memories trustworthiness. The local (memristor internal) temperature plays a key role in the resistive switching mechanisms [6]–[9]. The conductive filament that grows/dissolves within the memristor is affected by its temperature. Together with the consumed power, the circuit working temperature describes the dynamics of the conductive filament heating processes [10]. Consequently, characteristics such as switching speed [11] or data-retention ability [12] highly depend on the system working temperature. To the end of our knowledge, until [11] no driver had dealt with temperature caused problems.

But not only does the writing scheme and working conditions affect reliability. In the same manner, the device endurance strongly depends on the writing voltage characteristics [13], [14]. The use of a scheme able to guarantee the memory block operations does not ensure an optimal endurance for the devices and may cut back the useful lifetime. As presented in [15]–[17], depending on the application, reliable, power saving or faster profiles may be used. Therefore, usually a trade-off between these characteristics is fixed at design time [18].

Resistive switching based applications cover a wide range of scenarios. From non-volatile memories to reconfigurable devices, parallel datapaths to neural networks and approximated-computing processors, memristive devices will be required to work under varying conditions. If the working conditions vary -temperature range, computation load, etc.- the target application may benefit from a different and dynamically loaded profile. For example, aerospace circuits working at different altitudes experience environmental temperature changes. In such environments, RRAM data retention characteristics could improve by dynamically adapting the used writing amplitudes. The same principle could be applied to control the memory block endurance before employing drastic system level solutions relying on replicated hardware [19]. Taking advantage of a more suitable writing amplitude, related to a different HRS/LRS ratio [17], the memory lifetime would improve.

To address these issues, in this paper we present a writing driver adaptable to the memory working conditions that can dynamically load different writing profiles. We improve the driver presented in [11], supplying it with the additional hardware that allows achieving correct operations, no matter the working temperature, while focusing on performance, data retention, endurance or power consumption goals.

The circuit temperature catalyzes the switching mechanisms that define the device state. For low temperatures, higher voltages and or larger pulses (involving more power consumption) are required. Otherwise, using standard writing schemes would result in writing failures. Therefore, for critical applications

Fernando García-Redondo, Pablo Royer, Marisa López-Vallejo, Hernán Aparicio, Pablo Ituero and Carlos López-Barrio are with Technical University of Madrid - UPM, Spain. Corresponding author: fgarcia@die.upm.es.

This work was funded by the CICYT projects TOLERA TEC2012-31292 and TOLERA2 TEC2015-65902 of the Spanish Ministry of Economy and Competitiveness.

working at temperatures as low as 200 K, we must ensure the correctness of the operation. Similarly to *Dynamic Voltage Scaling* and *Frequency -DVS, DFS-* techniques, the dynamic driver adapts the RRAM memory writing scheme behavior to flexible targets, opening a wide range of programming strategies [15]–[17] to be dynamically applied. Thus, the driver selects among a myriad of writing strategies improving the system performance, power consumption, reliability, data retention abilities and HRS/LRS ratio durability, always considering the circuit working temperature as a critical parameter for the memory reliability.

The proposed architecture has been modularly designed, ensuring its compatibility with nowadays resistive-switching architectures. Thus, the dynamic driver works for both *1-Transistor-1-memristor (1T1R)*, *1-selector-1-memristor (1S1R)* and *1-memristor (1R)* configurations, built in standard *2D* or stacked *3D* architectures [20], [21].

In this work we first analyze the impact that circuit working temperature has on resistive memories reliability. We present the relationships between voltage amplitudes, pulse lengths and temperature, always guarantying the correct memory behavior. Using the model jointly developed by Stanford, Arizona and Beijing university [22], we will extract the writing requirements under which, at a specific temperature, the memory is able to be correctly written. This model takes into account the parasitic capacitance/resistance that appears together with the resistive switching structure.

In a second stage we present the proposed driver, describing the hardware modules that allow using dynamic profiles and detailing how it guaranties the circuit reliability while working on a wide range of temperatures.

In order to enlighten the driver capabilities, different writing profiles are analyzed:

- Profile sets with dissimilar HRS/LRS ratio, providing different device endurance [15]–[17].
- Profile sets offering extra protection against temperature misreads that may be caused by unexpected changes in the die thermal flow, or device to device thermal properties variations.

We bring into comparison their reliability capabilities and power consumption, and show the speed-ups this driver achieves when compared against other literature schemes.

The paper is structured as follows. First we introduce the memristor dependency on the temperature. Second, we address certain limitations that RRAM-based systems suffer. Third, the proposed driver is presented together with the novel dynamic profiles loading system. The full system implementation and related area cost is presented in Section V. Section VI presents the achieved speed-ups and power consumptions for different profiles. Finally, some concluding remarks are discussed in Section VI.

II. Resistive Switching Dependence on Temperature

The present paper uses the well known Verilog-A compact model of metal-oxide RRAM developed by Stanford University and Beijing University [22]. This model has been validated through several oxide devices and covers most physical variables/mechanisms involved in the resistive switching process. More importantly, this model emulates the physical device dependence on temperature [22], [23]. As seen in Figure 1, the device is composed of a conductive filament, located between the top and bottom electrodes, and surrounded by the Metal-Insulator-Metal (MIM) area. The device state will be between the low-resistance state (LRS) and the high-resistance state (HRS), depending on whether the CF is formed or dissolved. The transitions from HRS to LRS and LRS to HRS are called *SET* and *RESET* respectively. The used compact model describes the device using a conductive cylinder of variable height and width to emulate the CF dynamics.

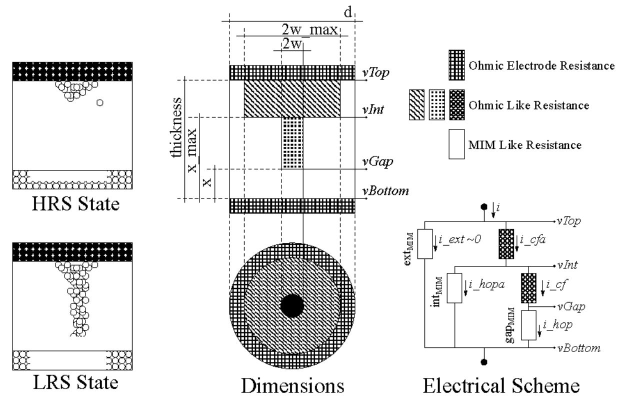

Fig. 1. HRS and LRS state representation together the used compact model.

The CF evolution is defined by the generation/recombination of oxygen vacancies and oxygen ions at the edge of the CF. For the *SET* operation, this model considers that the CF first grows in height until it reaches the opposite electrode, and then widens the cylinder base. Translated to the section scheme in Figure 1, first the gap x is reduced and then the width w would reach its maximum. On the other hand, the *RESET* process is modeled using x gap alone: the application of a negative voltage will result on the CF dissolution, and therefore on the increment of x until its maximum.

Both the applied voltage and CF temperature play key roles in CF dynamics [6]–[9], catalyzing the involved processes. A thorough study revealing the significant role of the circuit working temperature -and therefore the RRAM electrodes initial temperature- can be found in [24]. Being a_h, b_i, c_j, d_k constants, the speed of the CF evolution during the *SET* is determined by the following equations:

$$\frac{dx}{dt} = a_1 exp\left(\frac{-a_2 - a_3V}{T_{CF}}\right) \quad (1)$$

$$\frac{dw}{dt} = \frac{b_1}{w} exp\left(\frac{-b_2 - b_3V}{T_{CF}}\right) \quad (2)$$

while the *RESET* process speed is defined by

$$\frac{dx1}{dt} = c_1 exp\left(\frac{-c_2 - c_3V}{T_{CF}}\right) \quad (3)$$

$$\frac{dx2}{dt} = d_1 exp\left(\frac{-d_2}{T_{CF}}\right) sinh\left(\frac{d_3V}{T_{CF}}\right) \quad (4)$$

$$\frac{dx}{dt} = min\left(\frac{dx1}{dt}, \frac{dx2}{dt}\right) \quad (5)$$

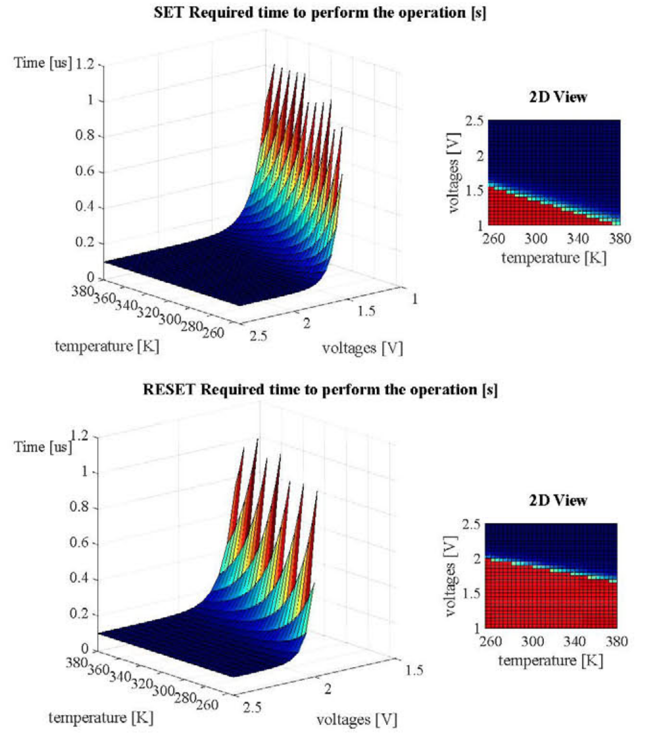

Fig. 2. Required time to perform standard SET and RESET operations using a fixed operation voltage, being T_0 the circuit working temperature. Operations where the CF did not achieve the desired geometry within the maximum considered -1.2 us-, and therefore requiring wider pulses, are plotted in red.

On the other hand, the current flowing through the device heats the CF. Naming R_{th} as the CF equivalent thermal resistance, the CF thermal behavior is modeled by the Joule-heating equation:

$$T_{CF} = T_0 + IVR_{th} \tag{6}$$

where T_0 refers the circuit working temperature.

The complex non-linear relation between the device state and the circuit operating temperature and applied voltage is displayed in Figure 2. As can be noted, the boundary conditions imposed by T_0 greatly influence the device dynamics. Both *SET* and *RESET* operations have been simulated at each pair of voltage amplitude and circuit temperature. The graphs represent the writing time required to completely perform the corresponding operation. Voltage amplitude-temperature pairs at which the operation is not correctly achieved have been computed but omitted to improve the graph readability. In agreement with [24], it can be concluded that, given a fixed operation length, to succeed in writing lower temperatures demand higher amplitudes. Therefore, and most especially in aerospace applications where working at extremely low temperatures is a requirement, the traditional two level amplitude scheme, one for *SET*, one for *RESET*, can no longer be applied.

III. Overcoming RRAM System Limitations

The required writing parameters map extracted from the analysis presented in the previous section, determines different parameter combinations that would lead to correct writing operations. If we wish to preserve the system performance using a fixed writing pulse length t_{op}, larger amplitudes are required for lower temperatures. From Figure 2 we could extract the minimum voltage amplitude, V_{abs}, able to achieve a correct writing operation within t_{op} s for all working temperatures within the considered range. Should there not be any non-desirable counter effects [25], the voltage amplitude V_{abs} could be employed satisfying the writing requirements in the whole temperature range. However, using V_{abs} as writing amplitude no matter the circuit temperature would raise two major disadvantages: not only would the power consumption be increased, but also the use of an amplitude larger than the required would lead to behavioral problems. This section gathers some of these concerning issues, all related to the RRAM device intrinsic nature.

A. Spureus Writing in Non-Selected Cells and Leakage

Memory blocks based on *1R* cells come with the advantage of not requiring any additional internal device to select the desired cell. However, given the *voltage-current* device behavior, applying the selection voltages to the chosen row and column lines would also apply a voltage difference in the cells located in these specific row and column, modifying their stored values. Consequently, voltages higher than the required aggravate spurious writings and may trigger the selector threshold. The widely used $V_w/2$ and $V_w/3$ writing schemes [4], [5], [21], [26] suffer from this spurious selection and expose cells to possible non-desired state alterations [22]. At the same time, this partial selection produces a leakage phenomena that should be addressed [27], [28]. More in depth information about how this stress affects data retention can be found in [12], [29], [30].

To solve this issue design solutions have been proposed by the addition of different elements to the *1R* cell scheme. A first solution introduces a transistor within the cell, building up *1T1R* cells. *1T1R* based schemes, as shown in [31], [32], introduce an additional path to access the transistor gate at each RRAM cell. In exchange of the additional logic, the design avoids accessing non-target cells, improving reliability. The inconvenient area increase coming with the additional CMOS transistor, together with supplementary *selection* line has lead to the investigation of alternative crossbar morphologies.

Most recent efforts focus on the implementation of *1S1R* cells, with the novelty of introducing an extremely non-linear device inside the cell, acting as a selector [33], [34]. Equivalent to two in series diodes, these innovative cells are intended to provide a customizable voltage threshold starting from which the cell would be selected.

B. Lifetime Reduction

The second non-desired effect, derived from the use of larger voltages, is thoroughly described in [13]. There, the device endurance (number of consecutive *SET-RESET* that a single device supports with no HRS/LRS ratio serious degradation) is studied and modeled. The results indicate that operation voltage amplitudes, and therefore switching speed, have a huge impact in the device degradation. The use of higher amplitudes would then reduce the RRAM lifetime. In a similar way, [35] relates endurance capabilities and the operation duty cycle, and therefore the applied energy. In [36] a different cycle-stress degradation model states the relation between HRS stress current and device degradation.

Similar conclusions can be found in recent works [14], [20], relating excessive operation voltage amplitudes with endurance reduction.

Regarding *1T1R* cells, the study accomplished in [15]–[17] indicated how different configurations of word-line (WL), bit-line (BL) and source-line (SL) voltages greatly affect the devices lifetime.

A myriad of studies have focused on accurately describing, modeling and palliating RRAM degradation. Thus, this stress-caused lifetime reduction has raised as one of the most concerning issues to be addressed before RRAM technologies get solidly established. Alternative schemes like [2] improve oxide devices endurance. The work in [19] evaluates the impact of RRAM read/write endurance on system failure, and even proposes reconfiguration approaches in which damaged resistive cells would be substituted by healthy ones. Even though this novel approach extends the system life time, it multiplies the consumed area.

C. Latency, Endurance, Power Consumption and Data Retention Trade-Offs

But not only does the applied writing voltage influence endurance. Write schemes that focus on data retention under stress conditions, may result in a faster device degradation, shortening its lifetime and imposing a hard trade-off [21], [37]. The solid conclusions presented in [15]–[17] enlighten the explicit trade-offs between HRS/LRS ratio, operation latency, device endurance, power consumption and data retention. Several writing configurations are proposed, emphasizing towards a specific application target.

D. Proposed Solution

When using RRAM based systems two major issues, regarding writing amplitudes and working temperature, need to be addressed. Firstly, higher voltages are needed to ensure the correct writing of memory blocks working at low temperatures. Secondly, the use of amplitudes larger than required would lead into functional errors, endurance and leakage problems.

We present an improved driver based on [11] to provide the most suitable writing amplitude depending on the working scenario and the application strategy. We propose a dynamic writing scheme that, similarly to *DVS* and *DFS* techniques, adapts the resistive switching blocks behavior to the circuit working temperature and application requirements. Consequently, our proposed solution increases the system adaptability and released the circuit designer from the hard trade-offs early imposed by a rigid, single amplitude driver.

IV. Dynamic Profile Thermal-Adaptive Write Architecture

The modular architecture of the proposed writing driver is represented in Figure 3. The basic structure presented in [11]

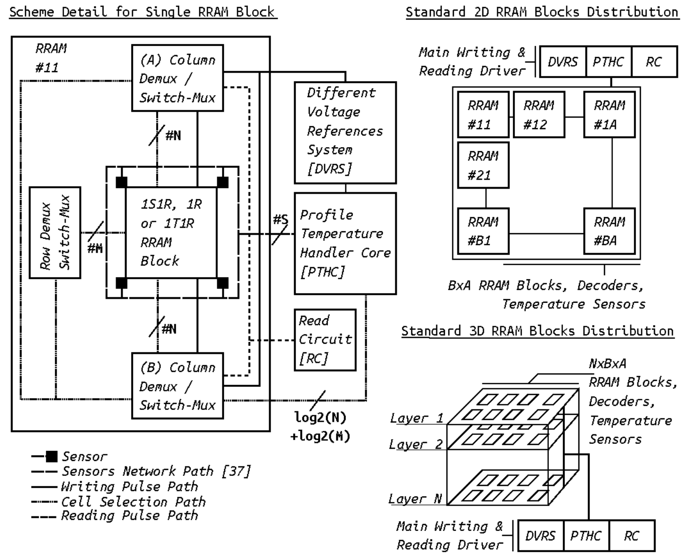

Fig. 3. Proposed writing architecture and interconnection of memory blocks in standard 2D and stacked 3D morphologies. The figure shows an S-sensors network connected to the temperature-handler core, together with the different voltage references supplier. The rows/columns demux-mux circuitry provides access to an $M \times N$ RRAM memory.

is improved, revising the maintained modules and introducing new ones required by new functionalities. Four main modules compose the driver:

- S temperature-sensors surrounding the memory blocks.
- A light communication network [38].
- The *profile-temperature-handler core.*
- The *different voltage references system.*

As described in $3D$ and $2D$ diagrams, the interconnection and distribution of surrounding sensors, submodules, and light-network guarantees design compatibility with both *1T1R* and *1R/1S1R* memory blocks: CMOS temperature sensors are placed next to the standard read and write access circuitry, close to the resistive memories. Low power CMOS temperature sensors present in the literature [39], or add-hoc sensor solutions can be interconnected using a light network [38]. This network scheme serves for calibration, sensing and transmission of any thermal information to the *profile-temperature-handler core*. Additionally, and thanks to its time-multiplexing scheme, this network could also drive supplementary information, such as devices aging [38], providing valuable intelligence to be used for the dynamic profile selection by the *profile-temperature-handler core*.

Using the proposed driver, all temperature-caused errors are fully mitigated, achieving the correct writing of the desired value. During this operation, a specific voltage amplitude is dynamically selected, ensuring that the writing process uses the most suitable voltage to accomplish the process regarding the loaded profile target.

A. Optimal Voltage Amplitude Selection

Writing operation starts by loading a temperature-voltage amplitudes profile in the *profile and temperature handler core*, targeting for a previously defined configuration (endurance, data retention, power consumption, etc.).

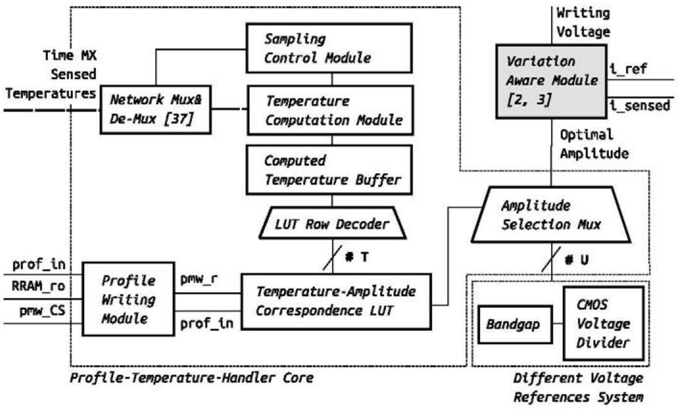

Fig. 4. Main core architecture interfacing complementary modules: *different voltage references system* and optional *variation aware module* [1], [40].

When a write operation starts, first, the *profile and temperature handler core* -or *PTHC*- decides whether a new temperature sensing should be performed or it relies on a previous measure. If a new temperature shall be acquired, a set of S different temperature measures are taken by the sensors surrounding the memory. These measurements are driven to the main core where the target RRAM cell temperature is computed. The required voltage amplitude is selected based on this temperature among a discrete set of values, and once driven to the selected cell, the writing operation takes place.

B. Profile-Temperature Handler Core

The *profile and temperature handler core*, represented by Figure 4, is in charge of the remaining operations.

First, the *sampling control module* manages whether or not a new sensing should take place. Based on the expected heat flow H_{max}, a sampling frequency f_s defines the error estimation [41]

$$\epsilon_{sampling} = \frac{H_{max}}{f_s}. \tag{7}$$

Usually, given the large inertia the temperature flow has, a low frequency is enough to correctly sense the memory temperature changes, barely affecting consumption and system performance. *Sampling control module* hardware implementation controls f_s relying on a simple configurable end-count counter.

If a new temperature measurement is commanded, *PTHC* first de-multiplexes the time-multiplexed thermal information broad-casted by the network, and reconstructs the temperature sensed by each one of the S sensors.

Using these temperature reads and the cell location, *temperature computation module* computes the cell temperature. This temperature value is stored in the *computed temperature buffer* to be used until -commanded by *sampling control module*- its value is overridden. This temperature value decodes the suitable writing voltage amplitude index stored in the look-up table -*LUT*-. Finally, this voltage index selects the desired amplitude, wired from the *different voltage references system* to the row/column drivers, achieving the satisfactory writing of the desired value in the target cell.

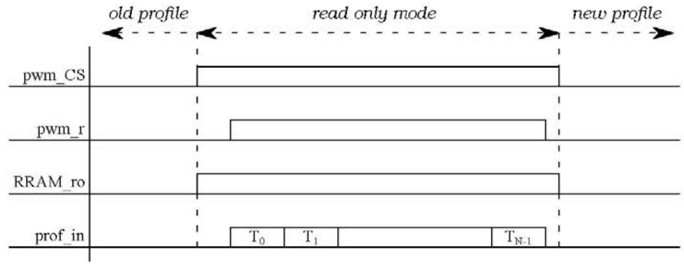

Fig. 5. Dynamic profile loading chronogram, *profile-writing-module*.

C. Dynamic Profiles

In [11] a new methodology for the calibration of physical devices is presented together with an algorithm for the selection of optimum voltage amplitudes regarding thermal-variations protection and power consumption.

As introduced in Section III, those voltages can be specified not only to ensure a correct operation no mater the working temperature, but to optimize different memory characteristics: device endurance, data retention, thermal-variations extra protection, power consumption or a solid HRS/LRS ratio. Different writing profiles suit different applications. Depending on the application orientation, a distinct profile would be desirable to match a particular target. Our proposal allows to dynamically change those oriented profiles, optimizing the memory behavior given a desired application target and/or environment characteristics.

The *profile writing module* subcircuit, shown in Figure 4, provides this functionality by safely replacing the values stored in the temperature-voltage amplitudes LUT for those related with the new profile to be loaded.

Profile Writing Process: The proposed scheme adds the hardware needed for the reconfiguration of the LUT locating the temperature-voltage decoding information. The *profile writing module* handles the process as detailed in Figure 5. To avoid possible writing problems, after the chip selector (`pwm_CS`) signal enables the profile-writing module, the resistive memory block is momentarily set to read-only mode. When the reconfigure signal, `pwm_r`, is activated, the new temperature-voltage decoding codes are in-series stored within the LUT. At the end of the process, and after `pwm_r` signal is deactivated, the resistive memory block is restored to write/read mode, and the updated profile can be used.

D. Different Voltage References System

The *different voltage references system* provides the required U discrete stable voltage amplitudes. To guarantee the stability of these signals, a bandgap voltage reference module has been used.

Figure 6 describes the implemented system. Instead of using U bandgap modules to provide U different outputs, we employed a single bandgap that stabilizes the voltage `V_BG`, followed by a CMOS voltage divider that generates the remaining amplitudes. This approach reduces the consumed area to near $1/U$ of the full version. Additionally and due to the `n`-transistor proximity, by using the CMOS voltage divider we reduce process variation effects. At design stage, by

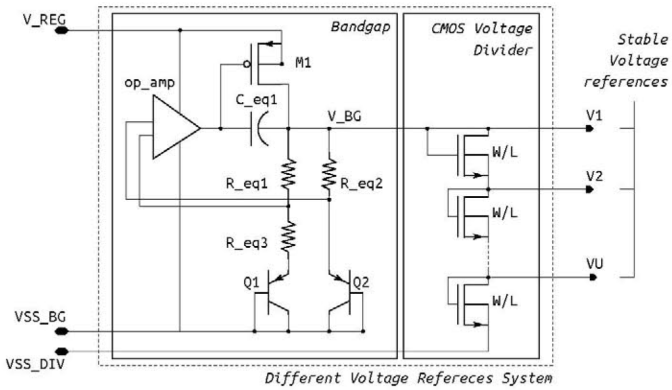

Fig. 6. *Dynamic voltage references system*, composed of a bandgap voltage reference followed by an U stage CMOS voltage divider.

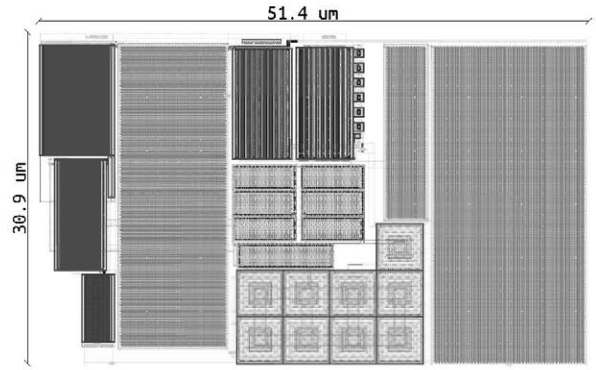

Fig. 7. Implemented 40 nm bandgap voltage reference.

tunning `R_eqX` and `C_eq1` equivalent resistors and capacitors the bandgap robustness against PVT can be optimized [42].

E. Optional Modules

Variability in RRAMs is one of the most concerning issues to be solved in memristive applications. Both device to device and cycle to cycle fluctuations can compromise the system reliability. To fight against variability caused problems, systems like [1], [3], [40] provide variability protection during write operations.

The modular design of the proposed driver allows embedding this kind of solutions. Figure 4 shows how the writing voltage leaving the *temperature handler core* would be controlled by the *variation aware module*. Using the reference `i_ref` and feedback `i_sensed` currents, the *variation aware module* controls the operation state and handles the writing voltage accordingly.

V. Full System Implementation

Together with *profile writing module*, the study of the hardware sizing and insights are thoroughly described.

To study the hardware sizing and area cost related to the proposed driver, different implementations have been built using commercial 40 nm CMOS technology.

A. Profile and Temperature Handler Core

1) LUT implementation: The area overhead introduced by the *profile-temperature-handler core* is mostly dominated by both LUT's and combinational hardware contributions. Different systems with dissimilar temperature resolutions have been designed and simulated. We extracted from these results that a temperature resolution of $2°C$ is enough to mitigate all thermal-variation caused writing errors. As the considered temperature range goes from 200 K to 380 K, using 7 bits to decode the selected amplitude, a 128 word LUT, with 4 bits per word was implemented. The use of a compact SRAM implementation only requires 14×63 μm^2. Those memories whose RRAM cells dependence on temperature requires from higher resolution, could easily increase the LUT size: for $1°C$ and $0.5°C$ resolutions, the LUT area would increase by $2\times$ and $4\times$ respectively.

2) Remaining Circuitry: The rest of the circuitry has been first described in *HDL* description language for its later synthesis using standard cells provided by the used 40 nm technology.

The functioning of the interconnection network [38] eases the scalability of the incorporated temperature sensors. As the information is time-multiplexed in the transmission, and the temperature computations are performed and buffered over fixed registers, the area cost of increasing the number of sensors only increases the word-length of some used registers. Thus, being 7 the bits used to encode the temperature, the cost of use $S = 2^R$ sensors would be translated into a $7 + R + 1$ registers word-length. Furthermore, as the *temperature computation module* requires a division during the temperature calculus, defining S as a power of two allows us to avoid a costly $1/S$ divider and use instead a simple register shifter, saving both power and area. The synthesized *profile and temperature handler core* without its LUT consumes a total area of 36.63×30.18 μm^2.

B. Different Voltage References System Implementation

Figure 7 shows the full custom compact design -only $51.4um \times 30.9um$- of the used bandgap voltage reference [42]. As introduced in the Section IV, due to the voltage divider regular design together with the n-transistor proximity, the voltage references system is hardened against process variations. For the experiments shown in next section the following voltage levels supplied the *different voltage reference system*: `V_REG`$= 3.3V$, `VSS_BG`$= 2.1V$, `VSS_DIV`$= 1V$. Using the above supply signals, all the required 17 voltage levels, from $2.6V$ to $1V$, were correctly generated.

VI. Temperature Aware, Reliable Oriented Profiles Simulation and Results

As introduced before, several works have thoroughly studied writing profiles focused on durability or power consumption. In this work we will analyze the power consumption and performance impact of nine independent profiles, resulting from the combination of three distinct pairs of $R_{ON} - R_{OFF}$ ratios and three different protection levels against temperature variation. Supposing each $R_{ON} - R_{OFF}$ ratio has a

TABLE I
CF CYLINDER SIZES RELATION WITH DEVICE'S RESISTANCE, USING A READ VOLTAGE OF $0.1\ V$. RELATIVE SIZES COMPARED AGAINST WIDEST (w_{MAX}) CYLINDER AND LARGER GAP (x_{MAX}) [22], BEING $3.65\ M\Omega$ AND $5\ K\Omega$ THE *HRS* AND *LRS* MODEL BOUNDARIES.

Profile	CF Relative Size	HRS	LRS	Ratio
P_{1X}	70%	$390\ K\Omega$	$10.26\ K\Omega$	38
P_{2X}	80%	$822\ K\Omega$	$7.86\ K\Omega$	104.6
P_{3X}	90%	$1.72\ M\Omega$	$6.21\ K\Omega$	276.9

TABLE II
DIFFERENT PROFILES AND MAXIMUM TEMPERATURE DEVIATION THEY CAN ASSUME AND MITIGATE.

Profile	MAX temperature deviation Δ_j
P_{X1}	$0°C$
P_{X2}	$5°C$
P_{X3}	$10°C$

characteristic durability [15]–[17], the user gains the ability of dynamically selecting the desired hardening level against temperature variations while choosing a more suitable configuration regarding the power consumption/endurance trade-off [15]–[17]. For each considered profile P_{ij}, i refers the *HRS-LRS* resistance pair and j the level of protection against thermal variations.

Section II described how the compact model emulates the device CF evolution using a conductive filament characterized by both magnitudes x and w. Table I shows the different *HRS-LRS* resistance levels used in this experiment and their relationship with individual combinations of the cylinder sizes.

We use the metrics s_{Pxy} and r_{Pxy} to measure the quality of the simulated operation. Using the computed sizes -x_{final}, w_{final}- of x and w CF magnitudes, and naming x_{Pxy}, w_{Pxy} as the CF sizes at which for a specific profile we consider the operation as correctly accomplished, the metrics can be defined as follows:

$$s_{Pxy} = \frac{max(0, w_{final} - w_{Pxy})}{w_{Pxy}} \times 100(\%) \tag{8}$$

$$r_{Pxy} = \frac{max(0, x_{final} - x_{Pxy})}{x_{Pxy}} \times 100(\%) \tag{9}$$

In those operations where s_{Pxy} or r_{Pxy} values are greater than zero the writing has not been correctly accomplished, and therefore a *writing error* arises. s_{Pxy} or r_{Pxy} values near 100% mean that the writing operation has not been able to store the desired value.

As introduced in Section V, the sensors and temperature evaluation scheme were sufficient to provide reliable temperature estimates and to mitigate all errors withing the $2°C$ resolution. However, embedded critical systems may require to decrease its working frequency -and therefore the temperature sampling frequency- which would lead to an increment of the computed temperature error. Additionally, due to aging effects the temperature sensor could increase its error. To address those scenarios, we consider profiles with additional protection against sensor malfunctioning errors or sudden and unexpected local temperature changes, which may cause errors not able to be mitigated. To that end we define three distinct P_{Xj} profiles with a different protection level against temperature variations. Therefore, as shown in Table II, using P_{Xj} the driver will mitigate all possible malfunctions caused by errors in the temperature estimation Δ_j.

Resistive switching speed is one of the concerns to be dealt with during design stages. In the following experiments we fix the writing operation to have a maximum length of $1\ \mu s$. To obtain the optimal voltage amplitudes that guarantee the correct operation in each considered temperature, we followed a specific RRAM cell characterization methodology. First, we simulated the updated RRAM model (*1R* and *1T1R* cell configurations) in `Cadence Spectre` simulator, using a $40\ nm$ commercial technology for CMOS components, and sweeping both circuit temperature and writing amplitude parameters. Given the capabilities extracted from the commercial CMOS *process design kit* models, for *1T1R* cells we considered a temperature range starting at $220\ K$. With those results we built 2-dimensional arrays containing the resulting values of x and w CF magnitudes.

We filter, using the previously defined metrics s_{Pxy} and r_{Pxy}, the circuit temperature-voltage amplitude pairs at which the writing operation is not correctly performed. Finally, we obtained the optimum temperature-voltage values matrices M_{ij} related to each profile P_{ij} after applying the voltage selector algorithm [11].

A. Required Voltages and Power Consumption Results

Tables I and II gather the characteristics of each one of the nine characterized profiles. Figure 8 summarizes the required voltages and consumed energies per operation, for both *1R* and *1T1R* cells depending on the configurable profile used by the driver. Higher HRS/LRS ratio profiles, related with data retention oriented strategies, use larger amplitudes and therefore consume more power than low HRS/LRS -and usually higher device endurance- profiles [17].

As expected, lower temperatures require higher voltages and energies. It can be noticed how in *1T1R* schemes, the in-series transistor modulates the current through the memristor, while diminishes the effective applied voltage. This effect raises two immediate consequences: First, for the same applied voltage *1T1R* cells are slower than *1R*. Second, slightly higher operation voltages are required.

For *SET* processes, the use of discrete voltage amplitudes and pulse lengths creates an increment in power consumption, producing the saw-tooth like curves, that can be compared against the most energy efficient curve. All required voltages can be coded using $4\ bits$ as the profiles consuming the larger number of different voltages, P_{3X} during *SET* process in *1T1R* cells, only required 11 levels. Therefore, the LUT area estimation described in Section V is enough to store all considered profiles. In the same way, the temperature-voltage amplitudes relation extracted from the initial study of Section II is presented for every simulated profile. For each profile, an almost decreasing linear staircase approximation can be used to model this voltage-temperature dependency. However, it

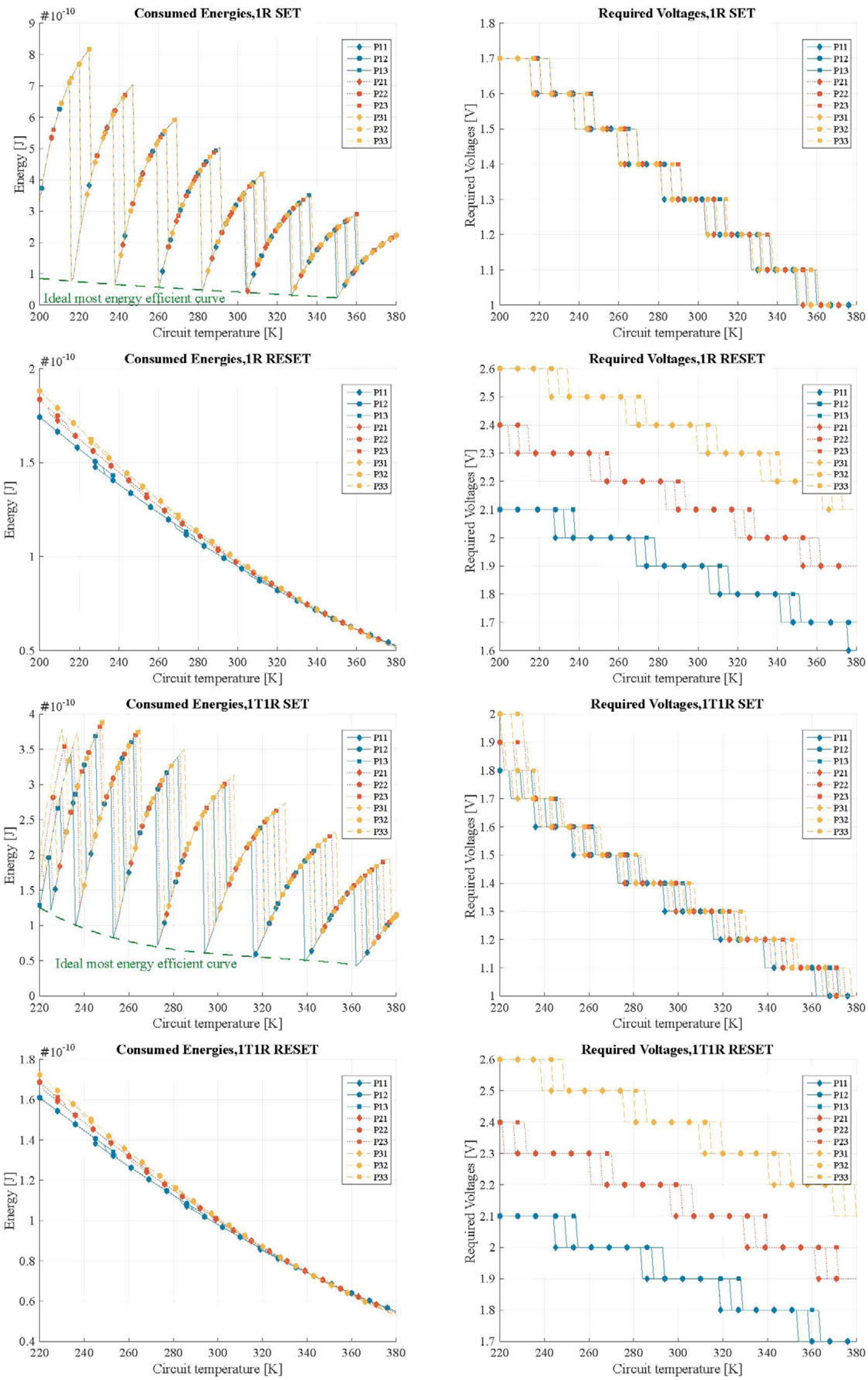

Fig. 8. Required voltages and consumed energies for every considered profile at each working temperature, for both *1R* and *1T1R* cells.

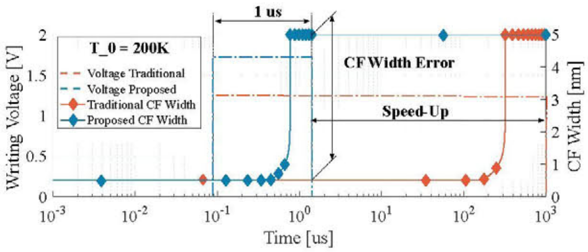

Fig. 9. *SET* operation comparison between proposed approach (in blue) and traditional one (in red) at $T_0 = 200K$. Traditional schemes would require from larger or consecutive writing pulses.

TABLE III
CALIBRATED VOLTAGE AMPLITUDES REQUIRED FOR EACH CONSIDERED PROFILE.

Cell Type	Operation	Required Amplitude		
		P_{1X}	P_{2X}	P_{3X}
1R	SET	1.25	1.25	1.3
	RESET	1.85	2.1	2.3
1T1R	SET	1.3	1.3	1.3
	RESET	1.9	2.1	2.35

must be remarked that, as the simulated temperature gets closer to the CMOS *process design kit* limit, this approximation can no longer be used. As shown in the *1T1R* cell *SET* process, lower temperatures exhibit an exponential shape.

B. Reliability and Speed-Up Results

Until [11], drivers presented in the literature varied the *SET/RESET* processes duration to correctly achieve the desired operation in presence of PVT variations using two distinct schemes. A first one enlarges/decreases the writing pulse until the operation is correctly accomplished [1], [40]. A second scheme consecutively performs write-read operations until the device state is satisfactorily switched [43], approach that can be complemented with incremental amplitudes [2]. Both protection mechanisms dramatically reduce the system performance in presence of high temperature variations. To correctly perform the switching at temperatures lower than the nominal (temperature at which the driver is calibrated, usually 300 K), both first and second variable pulse-length approaches would require from larger operation lengths. The transient displayed in Figure 9 compares a *SET* operation using the proposed approach and the traditional one.

We have computed the inaccuracy of using a scheme considering fixed voltage amplitudes, no matter the working temperature conditions, together with a maximum writing pulse of t_{op}. In the same manner, we have simulated the time required to perform *SET/RESET* processes for each P_{iX} dynamic profiles, comparing the operation speed when using both the proposed scheme and [40]. The first step in the procedure to simulate the writing approach [40] corresponds to the calibration of *SET* and *RESET* writing amplitudes. Using the fixed 1 μs pulse length we calculated the minimum amplitudes that completely accomplish the desired operation,

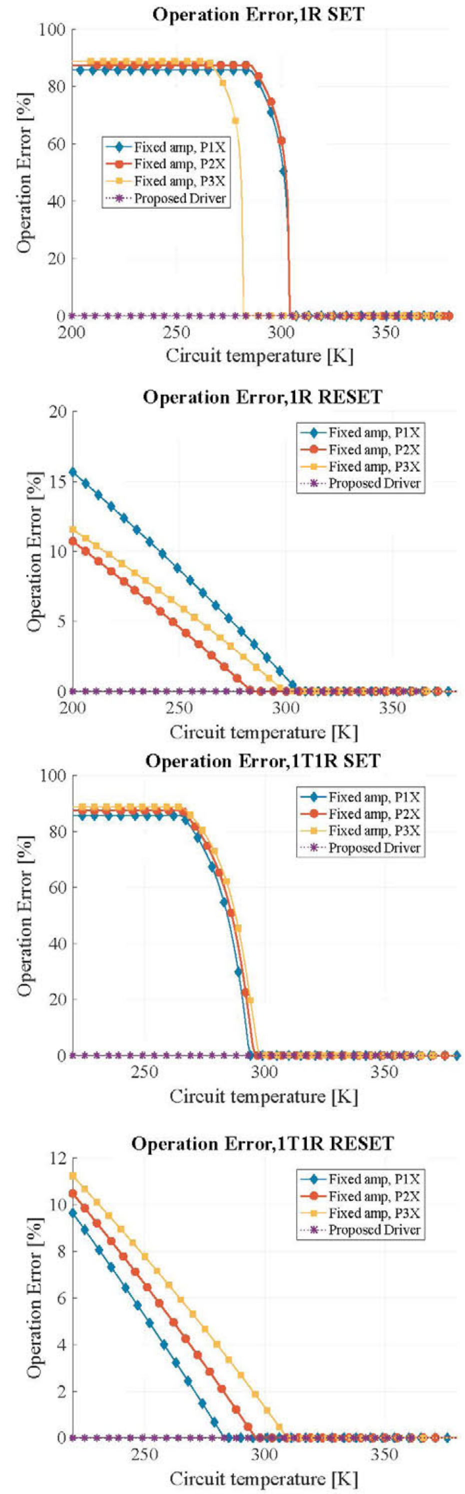

Fig. 10. Operation error dependence on temperature. While the scheme using the proposed driver is able to correctly behave under low temperatures, the classic fixed-voltage amplitude driver produces operation errors.

for each considered profile. Those amplitudes, captured in Table III, were used to compute the operation inaccuracy under fixed voltage pulse length and amplitudes, as well as the pulse length required to perform the device switching at each considered temperature, for the different profiles.

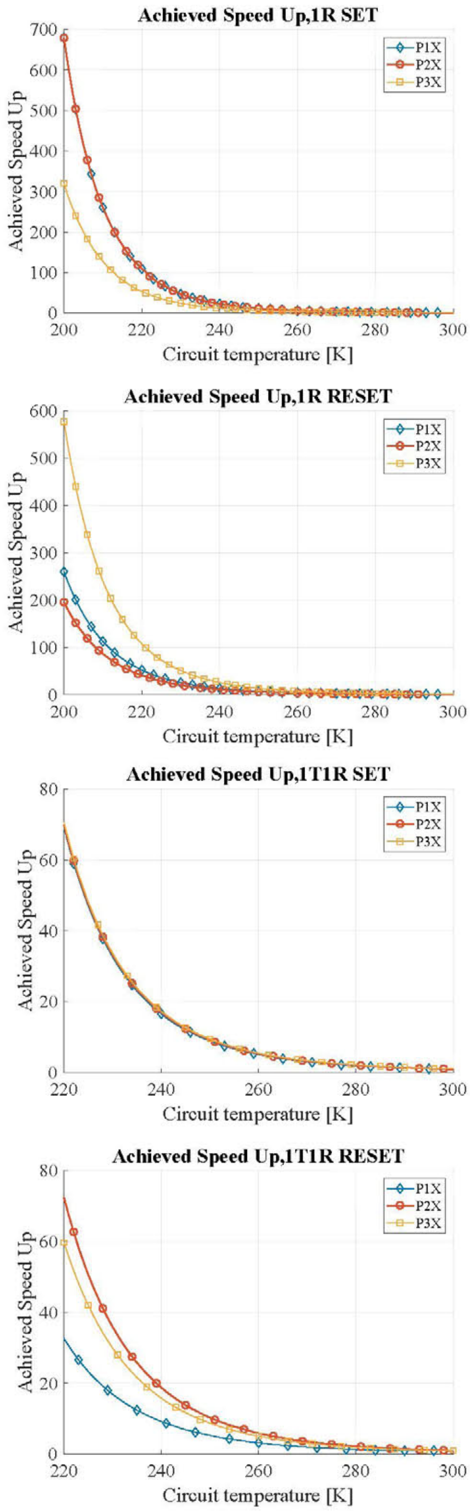

Fig. 11. Achieved speed-ups comparing the proposed scheme against variable-pulse-length schemes presented in literature under low-temperature conditions.

1) Fixed-Amplitude Scheme Error Results: To characterize how inaccurate a fixed voltage amplitude scheme can behave working under a fixed writing time, we simulated both $1R$ and $1T1R$ memories using 1 μs as the limiting operation length. Figure 10 shows the errors obtained after applying s_{Pxy} and r_{Pxy} metrics to the simulation results. It can be extracted that low temperatures produce a dramatic effect undermining the fixed-amplitude driver efficacy.

2) Speed-Up Results: We compare the minimum time required to correctly achieve a writing operation using a fixed-voltage amplitude scheme with variable pulse length against our fixed 1 μs, dynamic amplitude driver. To this purpose we computed the pulse length that satisfies both s_{Pxy} and r_{Pxy} metrics, ensuring 0% error at each considered temperature.

In our proposed architecture no matter the working temperature our driver guarantees the system 1 μs operation, while works like [2], [40], [43] would enlarge the operation length, or even repeat it until the cell achieves the desired state.

Given the cell dependence on the circuit working temperatures, for traditional drivers we find two different temperature sub-ranges, divided by the *room temperature* at which the driver writing voltages have been set.

For temperatures over room temperature, in schemes like [40], even though the device switching may occur earlier choking the cell voltage feeding, the writing operation timing would be guarded. Consequently, no timing difference would exist compared against our driver.

However, significant speed-ups -ratio between required time using the cited literature approaches and our proposed one- appear at lower temperatures. As shown in Figure 11, speed-ups near to 700× *-1R-* and 70× *-1T1R-* have been accomplished depending on the used profile, showing one the great advantages of using a variable amplitude scheme.

VII. Conclusions

Reliability, power consumption and device durability are the most pressing issues to be addressed to reinforce RRAM as the leading technology of future storage and novel computing paradigms. In particular the close relationship between writing process characteristics and RRAM features impose hard trade-offs during system design process. On the other hand, the circuit temperature strongly affects RRAM behavior, catalyzing the switching processes and, therefore, the memory block characteristics.

In [11] we proposed an innovative adaptive writing architecture for RRAM memories, able to correct any behavioral error that may appear during writing process due to temperature variations. In this paper we improve the driver introducing characteristics reconfiguration, allowing different writing approaches to be dynamically applied. By using the proposed solution and depending on the working scenario, less power, more durable or more reliability oriented profiles can be employed. Therefore, more efficient and optimized configurations can be utilized in critical applications such as aerospace circuits, while being able to ensure the correct behavior in a wide range of temperatures.

Nine different profiles, varying the cell HRS/LRS ratio and the level of protection against temperature sensing errors have been analyzed, characterizing their power consumption and required voltage amplitudes. Additionally, we have compared the proposed approach against other drivers able to handle RRAM thermal-caused miss-behaviors, obtaining notable writing operation speed-ups near to 700× in certain working temperatures.

REFERENCES

[1] A. Ghofrani *et al.*, "A Low-Power Variation-Aware Adaptive Write Scheme for Access-Transistor-Free Memristive Memory," *ACM J. Emerg. Technol. Comput. Syst.*, vol. 12, no. 1, pp. 1–18, aug 2015.

[2] K. M. Kim *et al.*, "Voltage divider effect for the improvement of variability and endurance of TaOx memristor," *Sci. Rep.*, vol. 6, no. October 2015, p. 20085, 2016.

[3] M.-F. Chang *et al.*, "Low VDDmin Swing-Sample-and-Couple Sense Amplifier and Energy-Efficient Self-Boost-Write-Termination Scheme for Embedded ReRAM Macros Against Resistance and Switch-Time Variations," *IEEE J. Solid-State Circuits*, vol. 50, no. 11, pp. 2786–2795, nov 2015.

[4] S. J. S.-J. Ham, H. S. H.-S. Mo, and K.-S. K. S. Min, "Low-power VDD/3 write scheme with inversion coding circuit for complementary memristor array," *IEEE Trans. Nanotechnol.*, vol. 12, no. 5, pp. 851–857, sep 2013.

[5] B. Govoreanu, L. Zhang, and M. Jurczak, "Selectors for high density crosspoint memory arrays: Design considerations, device implementations and some challenges ahead," in *2015 Int. Conf. IC Des. Technol.* IEEE, jun 2015, pp. 1–4.

[6] C. Walczyk *et al.*, "Impact of Temperature on the Resistive Switching Behavior of Embedded HfO2-Based RRAM Devices," *IEEE Trans. Electron Devices*, vol. 58, no. 9, pp. 3124–3131, 2011.

[7] C. Ahn *et al.*, "Temperature-dependent studies of the electrical properties and the conduction mechanism of HfOx-based RRAM," *Proc. Tech. Progr. - 2014 Int. Symp. VLSI Technol. Syst. Appl. VLSI-TSA 2014*, pp. 1–2, 2014.

[8] E. Yalon *et al.*, "Validation and Extension of Local Temperature Evaluation of Conductive Filaments in RRAM Devices," *IEEE Trans. Electron Devices*, vol. 62, no. 11, pp. 3671–3677, 2015.

[9] H. Sasakura, Y. Nishi, and T. Kimoto, "Temperature dependence of conductance in NiO-based resistive switching memory showing two modes in the forming process," *Appl. Phys. Lett.*, vol. 107, no. 23, p. 233510, dec 2015.

[10] H. Li *et al.*, "A SPICE model of resistive random access memory for large-scale memory array simulation," *IEEE Electron Device Lett.*, vol. 35, no. 2, pp. 211–213, 2014.

[11] F. Garcia-Redondo, M. Lopez-Vallejo, and P. Ituero, "A thermal adaptive scheme for reliable write operation on RRAM based architectures," in *2015 33rd IEEE Int. Conf. Comput. Des.* IEEE, oct 2015, pp. 367–374.

[12] N. Ciocchini and D. Ielmini, "Pulse-Induced Crystallization in Phase-Change Memories Under Set and Disturb Conditions," *IEEE Trans. Electron Devices*, vol. 62, no. 3, pp. 847–854, mar 2015.

[13] P. Huang *et al.*, "Analytic model of endurance degradation and its practical applications for operation scheme optimization in metal oxide based RRAM," *Tech. Dig. - Int. Electron Devices Meet. IEDM*, pp. 597–600, 2013.

[14] S. Balatti *et al.*, "Voltage-Controlled Cycling Endurance of HfOx-Based Resistive-Switching Memory," *IEEE Trans. Electron Devices*, vol. 62, no. 10, pp. 3365–3372, oct 2015.

[15] M. Mao *et al.*, "Programming strategies to improve energy efficiency and reliability of ReRAM memory systems," in *2015 IEEE Work. Signal Process. Syst.* IEEE, oct 2015, pp. 1–6.

[16] ——, "Optimizing latency, energy, and reliability of 1T1R ReRAM through appropriate voltage settings," in *2015 33rd IEEE Int. Conf. Comput. Des.*, vol. 85287. IEEE, oct 2015, pp. 359–366.

[17] ——, "Optimizing Latency, Energy, and Reliability of 1T1R ReRAM Through Cross-Layer Techniques," *IEEE J. Emerg. Sel. Top. Circuits Syst.*, pp. 1–12, 2016.

[18] I. Kazi *et al.*, "Energy/reliability trade-offs in low-voltage ReRAM-based non-volatile flip-flop design," *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 61, no. 11, pp. 3155–3164, 2014.

[19] peyman Pouyan, E. Amat, and A. Rubio, "Memristive Crossbar Memory Lifetime Evaluation and Reconfiguration Strategies," *IEEE Trans. Emerg. Top. Comput.*, vol. 6750, no. c, pp. 1–1, 2016.

[20] M. Yu *et al.*, "Novel Vertical 3D Structure of TaO x - based RRAM with Self-localized Switching Region by Sidewall Electrode Oxidation," *Nat. Publ. Gr.*, no. November 2015, pp. 1–10, 2016.

[21] P.-y. Chen, Z. Li, and S. Yu, "Design Tradeoffs of Vertical RRAM-Based 3-D Cross-Point Array," *IEEE Trans. Very Large Scale Integr. Syst.*, pp. 1–8, 2016.

[22] H. Li *et al.*, "Variation-Aware, Reliability-Emphasized Design and Optimization of RRAM Using SPICE Model," in *Des. Autom. Test Eur. Conf. Exhib. (DATE), 2015*, 2015, pp. 1425–1430.

[23] P. Huang *et al.*, "Parameters extraction on HfO¡inf¿X¡/inf¿ based RRAM," in *2014 44th Eur. Solid State Device Res. Conf.*, vol. 4, no. 7, 2014, pp. 250–253.

[24] C. Vaca *et al.*, "Study From Cryogenic to High Temperatures of the High- and Low-Resistance-State Currents of ReRAM Ni-HfO2-Si Capacitors," *IEEE Trans. Electron Devices*, pp. 1–7, 2016.

[25] A. Ghofrani, M. A. Lastras-Montano, and K. T. Cheng, "Toward large-scale access-transistor-free memristive crossbars," *20th Asia South Pacific Des. Autom. Conf. ASP-DAC 2015*, no. Mim, pp. 563–568, 2015.

[26] I. Vourkas *et al.*, "Alternative Architectures Toward Reliable Memristive Crossbar Memories," *IEEE Trans. Very Large Scale Integr. Syst.*, vol. 24, no. 1, pp. 206–217, jan 2016.

[27] P.-y. Chen and S. Yu, "Compact Modeling of RRAM Devices and Its Applications in 1T1R and 1S1R Array Design," *IEEE Trans. Electron Devices*, vol. 62, no. 12, pp. 4022–4028, dec 2015.

[28] A. Serb *et al.*, "Practical Determination of Individual Element Resistive States in Selectorless RRAM Arrays," *IEEE Trans. Circuits Syst. I Regul. Pap.*, pp. 1–9, 2015.

[29] Z. Fang *et al.*, "Resistive RAM Based on Hfo x and its Temperature Instability Study," *World Acad. Sci. Eng. Technol.*, vol. 48, no. 12, pp. 905–907, 2010.

[30] N. Raghavan *et al.*, "Statistics of retention failure in the low resistance state for hafnium oxide RRAM using a Kinetic Monte Carlo approach," *Microelectron. Reliab.*, pp. 2–6, jul 2015.

[31] C. Yakopcic *et al.*, "Generalized Memristive Device SPICE Model and its Application in Circuit Design," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 32, no. 8, pp. 1201–1214, aug 2013.

[32] W. G. Bennett *et al.*, "Dynamic Modeling of Radiation-Induced State Changes in HfO2/Hf 1T1R RRAM," *IEEE Trans. Nucl. Sci.*, vol. 61, no. 6, pp. 3497–3503, 2014.

[33] Sung Hyun Jo *et al.*, "3D-stackable crossbar resistive memory based on Field Assisted Superlinear Threshold (FAST) selector," in *2014 IEEE Int. Electron Devices Meet.*, no. 408. IEEE, dec 2014, pp. 6.7.1–6.7.4.

[34] S. Cortese *et al.*, "An amorphous titanium dioxide metal insulator metal selector device for resistive random access memory crossbar arrays with tunable voltage margin," *Appl. Phys. Lett.*, vol. 108, no. 3, p. 033505, jan 2016.

[35] H.-S. Choi, "Main degradation mechanism in AsTeGeSiN threshold switching devices," *Microelectron. Reliab.*, vol. 56, pp. 61–65, jan 2016.

[36] Y.-t. Chung *et al.*, "SET/RESET Cycling-Induced Trap Creation and SET-Disturb Failure Time Degradation in a Resistive-Switching Memory," *IEEE Trans. Electron Devices*, vol. 63, no. 6, pp. 2367–2373, jun 2016.

[37] M. Azzaz *et al.*, "Endurance/Retention Trade Off in HfOx and TaOx Based RRAM," in *2016 IEEE 8th Int. Mem. Work.*, no. c. IEEE, may 2016, pp. 1–4.

[38] P. Ituero *et al.*, "Light-Weight On-Chip Monitoring Network for Dynamic Adaptation and Calibration," *IEEE Sens. J.*, vol. 12, no. 6, pp. 1736–1745, jun 2012.

[39] P. Ituero, M. López-Vallejo, and C. López-Barrio, "A 0.0016 mm2 0.64 nJ leakage-based CMOS temperature sensor." *Sensors (Basel).*, vol. 13, no. 9, pp. 12 648–12 662, 2013.

[40] K. H. Jo *et al.*, "Self-adaptive write circuit for low-power and variation-tolerant memristors," *IEEE Trans. Nanotechnol.*, vol. 9, no. 6, pp. 675–678, 2010.

[41] P. Ituero, F. Garcia-Redondo, and M. Lopez-Vallejo, "Temperature sensor placement including routing overhead and sampling inaccuracies," in *2012 Int. Conf. Synth. Model. Anal. Simul. Methods Appl. to Circuit Des.* IEEE, sep 2012, pp. 69–72.

[42] F. García-Redondo *et al.*, "Characterization of Analog Modules: Reliability Analyses of Radiation, Temperature and Variations Effects," in *Des. Circuits Integr. Syst. Conf.*, 2016.

[43] F. Alibart *et al.*, "High precision tuning of state for memristive devices by adaptable variation-tolerant algorithm," *Nanotechnology*, vol. 23, no. 7, p. 075201, feb 2012.